

Práca s prirodzenými jazykmi za pomoci umelých nástrojov (prekladače, prekladové slovníky, prekladateľské pamäte, terminologické databázy a ich aplikácia v prekladateľskej praxi)

Mária Smetanová, Filozofická fakulta PU, majockas@gmail.com

Kľúčové slová: umelé spracovanie prirodzeného jazyka, elektronické zdroje, prekladače, prekladateľské pamäte, prekladateľské slovníky, terminologické databázy

Palabras clave: medios artificiales utilizados en la lengua, fuentes electrónicas, traductores automáticos, memoria traductora, diccionarios, base de datos terminológicos

Problematike rôznych typov prekladateľskej kompetencie je venovaných mnoho štúdií¹. Ich autori majoritne sústreďujú pozornosť na kompetenciu jazykovú (znalosť východiskového a cieľového jazyka), odbornú (ovládanie terminológie z tej-ktorej oblasti) či kulturologickú (prehľad v socio-kultúrnom pozadí). V našom príspevku sa zameriame na kompetenciu technickú (počítačová zdatnosť).

Súčasnosť nám prináša rad nových poznatkov a objavov na poli mnohých vedeckých disciplín. Progres, ktorý súvisí s globalizáciou sveta kladie na človeka veľké nároky a požiadavky, pretože ho doslova núti, aby sa s najnovšími výdobytkami doby zoznámil, prispôbil a plne zautomatizoval (v danom kontexte je na mieste zdôrazniť to, že rýchlosť pokroku nie je vždy úmerná aj rýchlosti ľudských schopností prispôbiť sa technickým vymoženostiam). Nové technológie prenikajú aj do profesie prekladateľa/tlmočníka a pokiaľ si chce udržať krok s modernou dobou je nútený trénovať svoje adaptačné schopnosti. Tendencia prepojiť translatologickú činnosť a svet technológií je evidentná. Gradujúca tenzia technického pokroku či trhu práce prispievajú k postupnej implementácii rôznych elektronických nástrojov aj do procesu prekladu. Využitie textových, grafických a tabuľkových súborov, elektronických slovníkov, terminologických databáz, nástrojov na automatický/strojový preklad, nástrojov na počítačom podporovaný preklad v prekladateľskom procese už prestáva byť raritou, ale stáva sa čírou realitou. Interakcia človeka a stroja je teda priam diktovaná a udržať krok s pretechnologizovanou dobou by malo byť pre prekladateľa výzvou, vecou zvýšenia jeho kompetencie. Na margo danej problematiky je potrebné zdôrazniť, že moderné informačné a komunikačné technológie sa stávajú pre človeka dobrým nástrojom len vtedy, ak ich vie vhodne použiť.

V príspevku je pertraktovaná téma umelých nástrojov (synonymá: elektronické nástroje, e-nástroje), ktoré sa používajú za účelom urýchlenia, zlepšenia a zefektívnenia práce prekladateľa. V centre pozornosti sú prekladače, prekladové slovníky, prekladateľské pamäte, terminologické databázy.

Nástroje na automatický/strojový preklad

V medzinárodnom kontexte je strojový/automatický preklad označovaný skratkou MT (angl. Machine Translation), príp. AT (angl. Automated Translation) či FAMT (Fully Automated Machine Translation). Strojový preklad² znamená počítačom zrealizovaný automatický preklad textu z jedného prirodzeného jazyka (východiskového jazyka) do druhého (cieľového), bez zásahu človeka do procesu prekladu. Ak hovoríme o nástrojoch na strojový preklad, máme na mysli rôzne prekladače, ktoré plnia funkciu prekladateľa, teda akoby ho „nahradzovali, zastupovali“. V našom príspevku sa budeme venovať internetovým prekladačom, ktoré pracujú v režime on-line, kedy je možné vložiť slovo, ale aj siahodlhý text do internetového prekladača a ten za veľmi krátky čas ponúkne jeho preklad. Je nevyhnutné

zdôrazniť, že počítačom ponúknutá verzia prekladu je iba pracovný preklad (tzv. podstročník), t. j. text, ktorý musí prekladateľ dôsledne skontrolovať, čiže dôraz sa kladie na jeho finálnu korektúru. Komplikácie s prekladom nastávajú hlavne pri polysémických výrazoch, homonymických jednotkách, reáliách, inotajoch, alúziách, ustálených slovných spojeniach, skratkách, slovách s preneseným významom, antroponymách, zaužívanej terminológii, tiež pri prevode typických gramatických kategórií, syntaktických konštrukcií atď., teda toho, čo je špecifické pre ten-ktorý prirodzený jazykový kód. Prekladač sa odporúča použiť pri preklade napríklad „bežných“ komunikačných fráz, „jednoduchších“ odborných textov. Výhodami prekladačov je ich rapidnosť, komfortnosť, siplicitnosť, dostupnosť. Aj napriek tomu, že v súčasnosti sú dosť sofistikované, ale žiaden z nich neponúka 100 % kvalitu, t. zn., že v mnohých prípadoch nezodpovedá originál prekladu. Z toho vyplýva, že prekladačom ponúknutá verzia má „ďaleko od dokonalosti“ a jej podobu je potrebné zrevidovať. Chceli by sme upozorniť na to, že tieto elektronické nástroje majú výsostne asistenčný charakter. Je potrebné mať na vedomí, že by sa mali používať až vtedy, keď prekladateľ dosiahne určitý kompetenčný stupeň východiskového a cieľového jazyka. Ľahko by sa mohlo stať, že neskúsený prekladateľ im natoľko podľahne, že sa bude spoliehať viac na nich ako na seba a neodhalí posuny, ktoré nastali pri strojovom preklade. Nedostatky či nezrovnalosti pri prevode strojovým prekladačom môžeme ilustrovať na niekoľkých exemplifikáciách. Google prekladač pri preklade zo slovenčiny do ruštiny náhodne vybraných, ale často používaných slov, slovných spojení, skratiek urobil nasledujúce chyby: národnosť (správne – национальность, nesprávne – гражданство), splnomocňujem (správne – удостоверяю, nesprávne – разрешаю), maturitné vysvedčenie (správne – свидетельство об экзамене на аттестат зрелости, nesprávne – аттестат), študijný odbor (správne – специализация, nesprávne – область исследования), štúdium na filozofickej fakulte (správne – учеба на философском факультете, nesprávne – исследование на факультете искусств), akademický titul (správne – степень, nesprávne – ученое звание), pracovná zmluva (správne – трудовой договор, nesprávne – рабочее соглашение), Výpis z registra trestov (správne – Справка о судимости, nesprávne – Сведения о судимости), technický preukaz (správne – технический паспорт, nesprávne – техническая лицензия), IČO, Z. z. (správne – ИНО, Св. з.; uvedené skratky prekladač nepreložil). Týmito príkladmi sme chceli poukázať na to, aké zradné môže byť použitie strojového prekladača.

Medzi významné české prekladače radíme APAC, Ruslan, Česílko. V súčasnosti sa populárnou stala paleta produktov českej spoločnosti Microton (Eurotran 2009, Wordmaster 2009 sú nástroje určené na preklad textových dokumentov na anglických, nemeckých webových stránkach do českého jazyka). K aktuálnym a často využívaným prekladačom patria PC Translator, Google Translate, SYSTRAN. V nasledujúcich riadkoch krátko predstavíme Google Translate, tiež SYSTRAN a zmienime sa aj o projekte EuroMatrix, ktorý je ešte v štádiu vývoja.

Google Translate (Google prekladač) je schopný prekladať z/do 58 jazykov (v štádiu testovania je ďalších 6 jazykových kódov). Je založený na princípe vkladania slov do prekladateľského systému a potom na báze rýchleho vyhľadávania identických, analogických slov, slovných spojení v paralelných textov na stránkach s obrovským množstvom dokumentov, ktoré už boli v minulosti niekým preložené a následnej analýze s vyhodnotením výskytu eventuálnych ekvivalentov systém vygeneruje preklad. V pracovnom prostredí Google prekladača je možné vidieť ako východiskový text, tak aj cieľový. Táto možnosť zrkadlového obrazu prekladateľovi výrazne uľahčuje finálnu kontrolu (korektúru). Google prekladač je on-line, t. j. voľne dostupný internetovým užívateľom. Medzi pracovnými jazykmi je aj slovenčina.

Spoločnosť SYSTRAN vyvinula niekoľko prekladateľských softvérov – SYSTRAN 7 Premium Translator, SYSTRAN 7 Business Translator, SYSTRAN 7 Office Translator,

SYSTRAN 7 Home Translator, ktoré podľa špecifikácie balíka toho-ktorého produktu môžu prekladať texty, e-maily, súbory v PDF formáte, prekladateľské pamäte, glosáre atď. On-line SYSTRAN prekladač, ktorý pracuje s 15 jazykovými kódmi je k dispozícii na <http://www.systranet.com/translate/>. Užívateľovi sú v pracovnom prostredí demonštrované dve okná, jedno pre východiskový text (označený ako Source), do ktorého sa vkladá originál a druhé pre cieľový text (označený ako Targe), kde prekladač ponúka jeho preklad. Medzi pracovnými jazykmi slovenčina nie je.

Na margo problematiky strojových prekladačov je potrebné spomenúť, že v súčasnej dobe Európska únia financuje vývoj nového prekladateľského systému *EuroMatrix*. Na projekte nového prekladača participuje ako EÚ, tak aj univerzity (napr. UK v Prahe) a inštitúcie (napr. česká prekladateľská agentúra CEET). *EuroMatrix* bude k dispozícii nielen orgánom EÚ, ale aj širokej verejnosti (internetovým užívateľom) už onedlho. Jeho použitie má byť bezplatné. Jazyková kapacita prekladača – 23 úradných jazykov EÚ, čiže 529 rôznych jazykových mutácií. Na rozdiel napríklad od konkurenčného prekladateľského systému Google bude *EuroMatrix* založený nielen na štatistickom modeli, lingvistickej analýze textov z jazykových korpusov, ale dôraz sa bude klásť aj na špecifiká toho-ktorého jazyka. Táto hybridnosť má garantovať čo najvyššiu kvalitu počítačom ponúkanej verzie prekladu.

Nástroje na počítačom podporovaný preklad

Počítačom podporovaný preklad poznáme pod skratkou CAT (angl. Computer-Aided Translation, Computer-Assisted Translation). Nástroje na počítačom podporovaný preklad sú založené na báze prekladateľskej pamäte (angl. Translation Memory). Reč ide o databázových programoch, ktoré profesionálni prekladatelia používajú s cieľom zachovania jednotnej terminológie a urýchlenia prekladateľského procesu. V žiadnom prípade sa nesmie zamieňať pojem strojový preklad (MT) s pojmom prekladateľská pamäť (TM), keďže rozdiel medzi nimi je práve v zásahu prekladateľa do procesu prekladu. MT spočíva v tom, že ho zrealizuje počítač, ktorý tým pádom nahrádza prácu prekladateľa v jeho prvotných fázach (prekladateľ vykonáva „už len“ finálnu korektorskú činnosť). Tvorcom TM je sám prekladateľ (jeho zásah je napr. pri ukladaní jednotlivých segmentov v priebehu prekladania do prekladateľskej pamäte). CAT nástroje (angl. CAT Tools) sa odporúčajú používať pri kvante textov určených na preklad, taktiež pri textoch odborných a textoch repetitívnych (pri opakovanom výskyte segmentu na základe zhody počítač vďaka sofistikovanému softvéru ponúka prekladateľské riešenia, ktoré sú zaznamenané v jeho pamäti). Práve to, že sa prekladové riešenia ukládajú do pamäte umožňuje ich opätovné použitie. Medzi populárne CAT nástroje patria softvéry SDL Trados, Wordfast, Transit, MetaTaxis, Dejavu, MemoQ, OmegaT, Ecross ai. V nasledujúcich riadkoch ponúkneme stručnú informáciu o softvéri SDL Trados Studio 2009.

SDL Trados[™], verzia SDL Trados Studio 2009 má difoltnú kapacitu niekoľko desiatok jazykových kódov. Tvorba jednotnej terminologickej banky zefektívňuje prácu prekladateľa. Softvér dáva možnosť práce v prostredí, kde je pre prekladateľa k dispozícii aj originál aj preklad. Jeho výhodou je, že neskresľuje textovú a grafickú štruktúru. Umožňuje dopĺňať do depozitára aj také dokumenty, ktoré neboli vypracované v Tradose. Taktiež dáva možnosť importovať a následne využívať pamäť iného prekladateľa. Slovenský jazyk je jedným z pracovných jazykov. Tento program patrí k proprietárnym softvérom, používanie ktorého je limitované kúpou licencie.

Elektronické prekladové slovníky

S nástupom počítačov prišli na trh aj elektronické slovníky, ktoré prekladateľa odbremeňujú od zdĺhavého hľadania hesiel v printových slovníkoch. Z pragmatických dôvodov mnohé z knižných slovníkov už majú aj svoju verziu na CD, príp. sú dostupné on-

line. V nasledujúcich riadkoch si predstavíme často využívané elektronické prekladové slovníky PC Traslator © 2010, Lingea slovníky a Logos.

PC Traslator © 2010 obsahuje 4 služby: slovník, prekladač textov, prekladač www stránok, prekladač označeného textu. V ponuke je 8 jazykových verzií. K dispozícii je na CD. Jeho mladšia verzia je k dispozícii on-line v 7 obojsmerných jazykových kombináciách na <http://slovník.azet.sk/>

Velký slovník Lingea Lexicon 5 predstavuje najnovší počítačový obojstranný slovník v 5 jazykových podobách.

Lingea dáva možnosť pracovať v režime on-line s prekladovými slovníkmi (24 možností výberu jazykovej kombinácie). Tieto elektronické jazykové nástroje sú dostupné na <http://slovníky.lingea.sk/>

Logos je multilingválny on-line slovník (prekladový, výkladový, synonymický, morfológický). Pracuje s niekoľkými desiatkami jazykových kódov. Je dostupný na <http://www.logos.it/>

Elektronické terminologické databázy

Správne pochopenie špecializovanej slovnej zásoby, ktorá sa používa v tej-ktorej oblasti je kľúčovým faktorom interlingválnej komunikácie. Disponovať terminologickými glosármi predstavuje pre prekladateľa odborných textov priam nevyhnutnosť. Neoceniteľným sa stáva, ak sú databázy termínov vypracované a dostupné, čo napomáha zefektívniť prekladateľskú prácu.

Pojem terminologická databáza zahŕňa súbor terminologických jednotiek z jedného alebo viacerých jazykov usporiadaných do terminologických záznamov podľa jednotlivých tematických okruhov. Predstavuje neoceniteľný zdroj informácií ako jazykového, tak aj odborného charakteru. V nasledujúcich riadkoch si priblížime databázu UNTERM a tiež databázu IATE.

Viacjazyčná terminologická databáza OSN (United Nation Multilingual Terminology Database), UN Database alebo *UNTERM* bola vytvorená s cieľom sprístupnenia nomenklatúry používanej nielen v OSN, ale aj mimo nej. V tomto elektronickom jazykovom nástroji sa nachádza niekoľko desiatok tisíc termínov. UNTERM dáva možnosť vyhľadávať terminológiu v 6 úradných jazykoch OSN. S touto bankou špeciálnej lexiky sa dá pracovať on-line na <http://unterm.un.org/>

Skratka *IATE* je odvodená od InterActive Terminology for Europe, čo v preklade znamená Európska interaktívna terminológia. Od r. 1999 sa prekladateľské útvary EÚ rozhodli pracovať na projekte vytvorenia jednotnej terminologickej databázy pre všetky inštitúcie EÚ. Od r. 2004 databázu začali využívať prekladateľské orgány EÚ a od r. 2007 bola sprístupnená širokej verejnosti (internetovým užívateľom). Táto viacjazyčná, interaktívna, inštitucionálna databáza termínov je denne dopĺňaná o nové pojmy a v súčasnosti obsahuje cca 8 mil. termínov. Terminologická banka je dostupná vo všetkých 23 úradných jazykoch členských krajín EÚ, tiež je doplnená aj o latinský jazyk, t. zn., že disponuje 24 jazykovými kódmi. Webová služba IATE je k dispozícii na <http://iate.europa.eu/>

Výhody a nevýhody umelých nástrojov (elektronických nástrojov, e-nástrojov):

- 1) vstupné náklady na ich zabezpečenie
— prvotné výdavky prekladateľa na to, aby disponoval týmito elektronickými nástrojmi sú pomerne vysoké (týka sa to nákupu licencií rôznych CAT nástrojov či slovníkov);
- 2) náročnosť s ich obsluhou
— mnohých prekladateľov odrádza od ich využitia „zložitost“ či „zdlhavost“ naučiť sa s nimi pracovať (máme na mysli prácu s CAT nástrojmi);
- 3) závislosť od elektronických nástrojov

— prekladateľ by sa v žiadnom prípade nemal stať ich „otrokom“, práve naopak, e-nástroje by sa mali stať jeho sluhami;

4) štandard v dnešnej dobe

— prekladateľ by mal držať krok s dobou, keďže čoraz častejšie môže byť jeho práca konfrontovaná s výdobytkami doby (zvýšené nároky, ktoré kladie trh práce, napr. niektoré prekladateľské agentúry vyžadujú znalosť práce s CAT nástrojmi);

5) nedôvera voči neznámemu, pre mnohých zbytočnému

— veľa prekladateľov s dlhoročnou praxou „nemá potrebu“ inklinovať k tomu, aby siahli po vymoženostiach súčasnej doby;

6) editor, korektor, kontrolór viac ako prekladateľ

— pri práci s rôznymi e-nástrojmi môže mať prekladateľ pocit, že veľa času venuje korektorskej činnosti ako samotnej prekladateľskej;

7) asistenčný (podporný, pomocný) charakter

— prekladateľ sa napr. v prípade strojového prekladu nemôže 100 % spoliehať na kvalitu výsledkov.

V našom príspevku sme pertraktovali tému umelých nástrojov (elektronických nástrojov, e-nástrojov). Zamerali sme sa na problematiku prekladačov, prekladových slovníkov, prekladateľských pamätí, terminologických databáz, aplikáciou ktorých sa má dosiahnuť akcelerácia, zlepšenie a zefektívnenie prekladateľského výkonu. Taktiež sme na základe komparácie zosumarizovali ich výhody a nevýhody. Naším cieľom bolo poskytnutie možnosti nazrieť do elektronického sveta, v ktorom už mnohí prekladatelia pracujú, keďže dichotómia – umelé nástroje (strojové prekladače, prekladateľské pamäte, prekladové slovníky, terminologické databázy) a ľudská kontrola – aplikované pri preklade postupne prestávajú byť hudbou budúcnosti prekladateľského odboru a stávajú sa bežnou praxou.

Poznámky

¹ Břejlová, D.: Osobnost konferenčního tlumočníka a profesionálního překladatele. Praha: FF UK v Praze 2001, 116 s.

Stahl, J.: Osobnost a vlastnosti tlmočníka. In: ToP, roč. XII, č. 56, 2001 s. 12

Vychodilová, Z.: Interkultúrna kompetence překladatele jako podmínka kvality překladu. In: Hľadanie ekvivalentnosti IV. Prešov: FF PU v Prešove 2009, s. 635 – 642.

Zázrivcová, M.: Prekladateľ – bytosť kritická a pochybujúca. In: <http://www.jtunion.org/spip/recherche.php3?recherche=Z%C3%A1zrivcov%C3%A1%2C+M.%3A+Prekladate%C4%BE+%E2%80%93+bytos%C5%A5+kritick%C3%A1+a+pochybuj%C3%BAca>

Zborník príspevkov „Kvalita prekladu a prekladateľské kompetencie“ z konferencie konanej 17. – 19. septembra 2003 v Budmericiach.

² V danom kontexte dochádza k prepojeniu lingvistiky a informatiky, teda hovoríme o počítačovej lingvistike, tzv. komputačnej lingvistike, ktorá je na pomedzí jazyka a umelej inteligencie. V tomto špecializovanom smere sú aplikované poznatky ako z lingvistiky, tak aj z výpočtovej techniky (umelé spracovanie prirodzeného jazyka).

Literatúra

BAKER, M.: Routledge Encyclopedia of Translation Studies. New York: Routledge 2001. ISBN 0415255171.

BERBER, D.: ICT (Information and Communication Technologies) in Conference Interpreting: a survey of their usage in professional and educational settings. In:

<http://www.google.sk/search?hl=sk&source=hp&q=Berber%2C+D.%3A+ICT+%28Informati on+and+Communication+Technologies%29+in+Conference+Interpreting%3A+a+surey+of+t>

heir+usage+in+professional+and+educational+settings.+&aq=f&aqi=&aql=&oq= [11. 01. 2011].

BOWKER, L.: Computer-Aided Translation Technology: Practical Introduction. Ottawa: University of Ottawa Press 2002.

BUCHTEL M. – PALATIN, H.: Bezplatné terminologické databázy. In: <http://www.productivetranslator.com/jeronymovy-dny-2010> [11. 01. 2011].

ČONDAK, M.: Strojový překlad integrovaný do CAT. In: <http://www.condak.net/akce/jeronym10/cs/00.html> [11. 01. 2011].

GARCIA, I.: Is machine translation ready yet? In: International Journal on Translation Studies. Target 22 (1), 2010, s. 7 – 21.

HAJIČOVÁ, E.: Lingvistické aspekty umělé inteligence. In: http://wiki.matfyz.cz/wiki/Lingvistick%C3%A9_aspekty_um%C4%9B%C3%A9_inteligence [11. 01. 2011].

HUTCHINS, W. J.: Machine translation: past, present, future. Chichester: Ellis Horwood 1986.

MAČURA, M.: Informačné a komunikačné technológie v univerzitnom jazykovom vzdelávaní pri vyučovaní prekladu. In: http://ec.europa.eu/slovensko/documents/archive/konferencia_02122009/dr._macura.doc [11. 01. 2011].

MARIOTHOVÁ, P.: Průručka k vyhledávání překladatelských řešení na internete – alebo Ako na to? In: <http://www.scribd.com/doc/53929947/Prirucka-k-vyhledavaniu-prekladatelsk%C3%BDch-rie%C5%A1eni-na-internete> [11. 01. 2011].

ROBINSON, D.: Becoming a Translator: An Introduction to the Theory and Practice of Translation. New York: Routledge 2003. ISBN 0415300339.

SELESKOVITCH, D. – LEDERER, M.: Pédagogie raisonnée de l'interprétation. Paris: Office des publications officielles des Communautés européennes 2002. s 361 – 371.

SOLOVJEVA, A. V.: Professional'nyj perevod s pomoščju kompjutera. Sankt-Peterburg: Piter 2008.

WILLS, W.: Knowledge and Skills in Translator Behavior. Amsterdam: Benjamins Publishing Company 1996. ISBN 90 272 1615 0.

http://translate.google.sk/about/intl/sk_ALL/ [11. 01. 2011].

<http://unterm.un.org/dgaacs/unterm.nsf/AboutUnterm?OpenPage> [11. 01. 2011].

<http://www.cdt.europa.eu/SK/whatwedo/Pages/IATE.aspx> [11. 01. 2011].

<http://www.ceet.cz/cat-preklady/> [11. 01. 2011].

<http://www.euromatrix.net/> [11. 01. 2011].

http://www.logos.it/about_logos/ [11. 01. 2011].

<http://www.microton.cz/> [11. 01. 2011].

<http://www.protranslator.eu/> [11. 01. 2011].

<http://www.systran.co.uk/translation-products/desktop> [11. 01. 2011].

<http://www.trados.com/en/> [11. 01. 2011].

<http://www.translationzone.com/en/> [11. 01. 2011].

Resumen

El presente artículo viene dedicado al tema de los medios artificiales (herramientas electrónicas) que se utilizan con el fin de efectivizar el trabajo de un traductor. Más detalladamente hablamos de los traductores automáticos, memorias traductoras, diccionarios y bases de datos terminológicos. Nos centramos en su característica, sus concreciones, ventajas y desventajas de su uso. El objetivo de este breve análisis es acercar el mundo electrónico con el que muchos traductores experimentan. Debido a que la dicotomía entre las herramientas artificiales y el control humano aplicado en el momento de realizar una traducción han dejado

de ser un mundo ficticio y están pasando a ser el “pan de cada día“ en la vida práctica de un traductor.